RESULTS FOR DEVELOPMENT INSTITUTE

Working Paper | May 2012

An exploratory application of data envelopment analysis to the efficiency of health service coverage and access

Rodrigo Moreno-Serra and Peter Smith
Business School & Centre for Health Policy, Imperial College London, United Kingdom

This paper was commissioned by the Results for Development Institute for the Transitions in Health Financing research initiative which is part of the Rockefeller Foundation's strategy on Transforming Health Systems. The goal of this research initiative is to foster greater understanding of the factors influencing health spending in low- and middle-income countries in an effort to identify and assess policies for achieving universal health coverage.

An exploratory application of data envelopment analysis to the efficiency of health service coverage and access

By Rodrigo Moreno-Serra[a] and Peter Smith[b]
Business School & Centre for Health Policy, Imperial College London, United Kingdom

May 2012

Prepared for the *Transitions in Health Financing* project

Acknowledgements

This paper was prepared for the Results for Development Institute as part of a project entitled "Transitions in Health Financing" and financed by the Rockefeller Foundation. The authors are grateful for comments on earlier drafts provided by Bill Savedoff, David de Ferranti, Dean Jamison, and participants in a review meeting held in Bellagio, Italy in May 2011.

Contact information:

[a] Email: r.moreno-serra@imperial.ac.uk
[b] Email: peter.smith@imperial.ac.uk

A Working Paper of the Results for Development Institute (R4D)
www.resultsfordevelopment.org

Introduction

Progress towards universal health coverage (UHC) entails securing access to needed health services for the population as a whole. A fundamental requirement for such progress is the removal of financial barriers to access, principally direct patient charges for the use of the services. This requires the creation of alternative 'pooled' sources of funding, in most countries taking the form of health insurance arrangements, funded by governments, employers or donor agencies. Yet the creation of such pooled funds does not necessarily lead to improvements in coverage. The purpose of this paper is to examine variations in the efficiency with which pooled funds are used to secure coverage, in the form of access to needed health services. The intention is to identify countries that – other things equal – appear to secure given levels of coverage with the lowest levels of pooled funding, to explore whether there are more general lessons that can be learned from these benchmark countries.

Inefficiency is intrinsically difficult to measure, as it represents the shortfall in performance from what could in principle be achieved, a concept that is manifestly open to challenge. The usual approach to inferring inefficiency has been to construct an estimate of the 'health production frontier' (or its analogue the cost function) on the basis of the observed performance amongst exemplar units of observation, in this case health systems (Jacobs, Smith and Street 2006). This was the principle underlying the *World Health Report 2000* and most of the subsequent efforts to assess health system performance (World Health Organization, 2000). The report of the WHR2000 Scientific Peer Review Group (SPRG) sets out the very many challenges associated with such an undertaking (Anand *et al* 2002).

Two approaches have dominated the productivity literature: econometric methods, pre-eminently various forms of statistical methods such as panel data models and stochastic frontier analysis (SFA), and the descriptive methods known as data envelopment analysis (DEA). Although these methods approach the task in radically different fashions, they have the common intention of using the observed behaviour of all organizations to infer the maximum feasible level of attainment (the production function), and offering estimates of the extent to which each individual organization falls short of that optimum.

The problems associated with using statistical models to infer health system efficiency have been exhaustively documented by the SPRG. They do not appear to offer much help in tackling questions underlying health system efficiency, because they require specification of a functional form, the nature of which is highly contested. If we then detect a systematic relationship between (say) UHC and efficiency, it will not be clear whether the result is due to a genuine association or to an incorrectly specified functional form (Smith and Street, 2005). In this report we therefore experiment with the use of non-parametric DEA and its derivatives to explore the link between efficiency and certain health system characteristics. We must emphasize that the analysis is exploratory and intended to illustrate the potential of the techniques rather than to produce definitive results. It is infeasible to present full results. Instead this report describes the scope of analysis. It is accompanied by a spreadsheet of results.

In this report we first outline the standard cross-sectional and dynamic uses of DEA. We follow this with an application of these methods to national performance on progress towards universal health coverage. We conclude with a brief discussion of the results and methods.

Data Envelopment Analysis

Data envelopment analysis is based on the economic principles of cost and production functions, but uses estimation techniques based on linear programming models. In summary, it searches for the organizations that 'envelope' all other organizations on the basis of a composite estimate of efficiency. For each organization, it looks for all other organizations that secure the same (or better) outputs at lowest use of inputs. Or conversely it can be used to search for the other organizations that use the same (or lower) inputs and secure the highest level of outputs. For each organization, the ratio of actual to optimal performance is referred to as inefficiency.

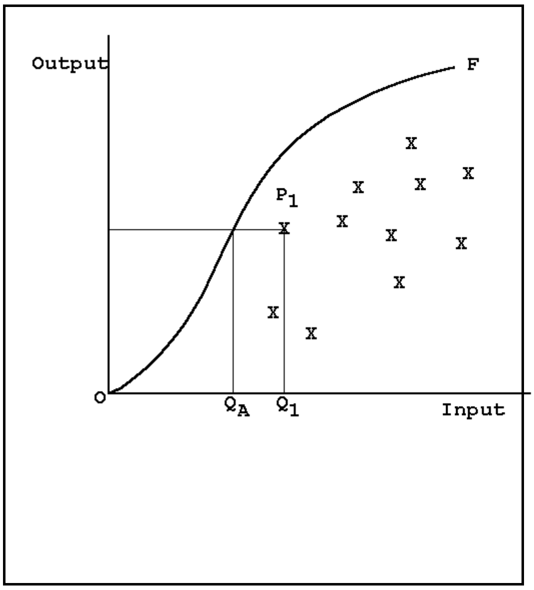

Figure 1: Production frontier, one input one output

The economic principles

Suppose there are a number of directly comparable organizations producing a single output, and that only one input is required to produce the output. Then we might observe a situation as shown in Figure 1. The curve OF represents the production frontier, showing, for a given level of input, the maximum output that is technically feasible. All organizations (shown as crosses) must therefore in practice lie on or below this curve.

Clearly organization P_1 is not 100% efficient. At its chosen level of output, it should be possible for P_1 to reduce its inputs from OQ_1 to OQ_A. In his seminal paper, therefore, Farrell (1957) deems the ratio $E_1 = OQ_A/OQ_1$ to be a reasonable measure of technical efficiency. It takes no account of price data, so there is no way of telling if the chosen level of output is optimal in an allocative sense. The measure E_1 always lies between zero (no output) and one (total efficiency).

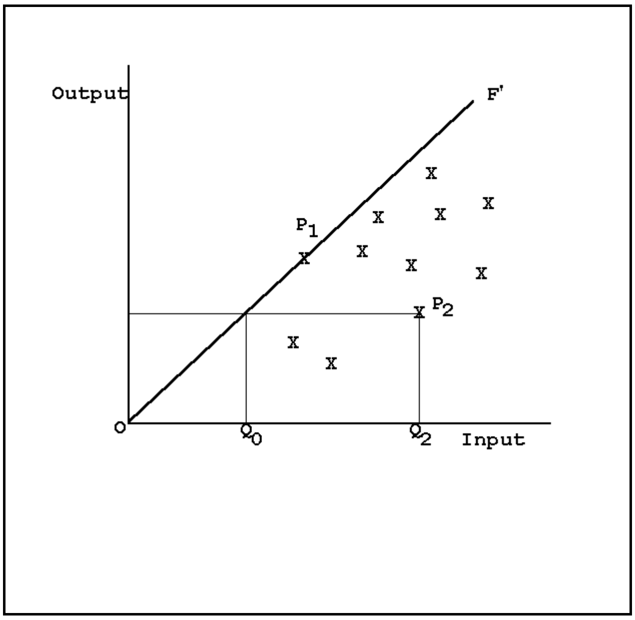

Figure 2: Linear production frontier, one input one output

In practice the nature of the curve OF is rarely known. Practical considerations therefore demand some simplifying assumptions. The obvious first one is:

Assumption 1: There are constant returns to scale. That is, the curve OF is a straight line through the origin.

This means that we can easily identify the most efficient organization: it is that with the highest ratio of output to input. This is ratio analysis at its simplest. Figure 2 reproduces Figure 1 with the theoretical production frontier replaced by a linear production frontier OF'. The problem of identifying the slope of OF' has been solved by employing:

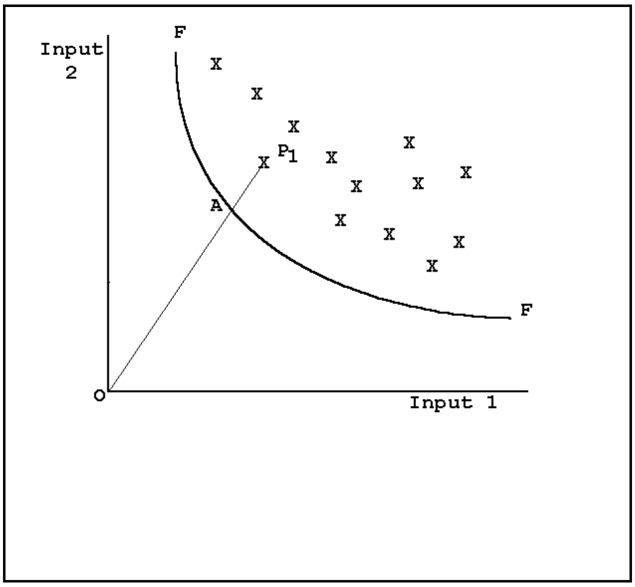

Figure 3: Production isoquant, two inputs

Assumption 2: The production frontier is defined by the most efficient organization. That is, there is always at least one efficient organization that defines the frontier, all inefficient organizations lying below the frontier.

By this criterion, P_1 is now efficient. The technical efficiency of the inefficient organization P_2 is measured by the ratio $E_2 = OQ_0/OQ_2$, where the point Q_0 has been determined not by some 'ideal' efficiency (as in Figure 1), but by the performance of another organization P_1 which appears to make better use of resources. In general, there is no way of knowing whether this empirical measure of efficiency overestimates or underestimates the true technical efficiency of an organization; that is, the measure of efficiency we should obtain if the curve OF were known.

Figure 3 illustrates the case with two inputs, but still a single output. Because of the constant returns to scale assumption, we can represent in two dimensions the inputs required by each organization to produce a given level of output. All organizations therefore lie on the production frontier FF (if they are efficient) or above it (if they are inefficient). The curve FF is the usual isoquant of economic theory. A measure of the technical efficiency of P_1 analogous to the one input case is then OA/OP_1. This shows the extent to which each of the inputs could be reduced if P_1 were efficient, whilst retaining the same mix of inputs. Notice that the ratio is independent of the scales on which the inputs are measured.

To make this notion empirically useful, the curve FF must once again be approximated with reference to the observed performance of organizations. Farrell's solution was to make the most conservative estimate of the frontier. This requires:

Assumption 3: The production frontier is convex to the origin, and has nowhere a positive slope. That is, along the frontier, reduced use of one input necessitates an increase (or certainly no decrease) in the use of the other input in order to maintain production.

Then the estimated frontier for the situation in Figure 3 is defined by the curve F'F' in Figure 4. This has the property that no segment has a positive slope, and no organization lies between it and the origin.

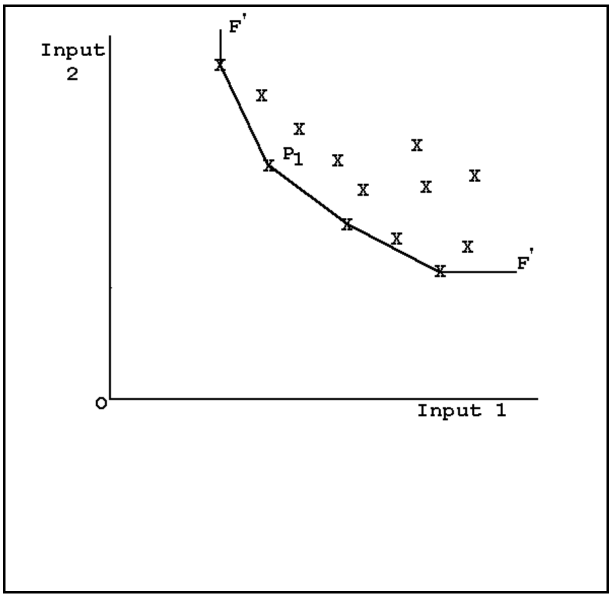

Figure 4: Piecewise linear isoquant, two inputs

The curve extends to infinity from the points with the lowest use of input per unit of output. Efficiency is now measured relative to this estimated isoquant, and P_1 is deemed efficient.

Figure 4 shows the estimated isoquant for just one level of output. Because of the assumption of constant returns to scale, the complete production frontier (for all levels of output) is easily inferred. By including output as a third dimension, it is possible to envisage the production frontier as a series of planes, extending from the origin, and passing through the line segments making up F'F'. All observed organizations lie on or within this 'envelope' of planes. This argument can be readily extended to any number of inputs and a single output.

Farrell's method lay neglected for many years. However, a paper by Charnes, Cooper and Rhodes stimulated fresh interest in the approach. One development was the ability to handle multiple outputs as well as inputs, which turns out to be computationally (if not conceptually) relatively straightforward using mathematical programming techniques. If there are m inputs and s outputs, then the production frontier becomes a surface in $m+s$ dimensional space. The efficiency of an organization is determined by the maximum distance it lies from that efficient surface - that is, the maximum extent to which it could improve all of its outputs - given that its existing level of inputs cannot be increased. Under the constant returns to scale assumption, this is equivalent to asking how much the organization's inputs could be reduced while maintaining existing output. The full mathematical formulation of the technique is known as data envelopment analysis (Charnes, Cooper and Rhodes 1978).

Since then the theory and practice of DEA has developed enormously (Emrouznejad *et al* 2008), although the basic principles remain unchanged. Compared to statistical methods, DEA has some attractive features. It requires none of the restrictive assumptions required to undertake regression methods. It can handle multiple inputs and multiple outputs simultaneously, and it requires none of the stringent model testing that is required of statistical techniques. Furthermore, if the interest lies in multiple criteria (i.e. outputs) to assess the performance of organizations, DEA does not require the analyst to pre-define weights for these criteria – yet the method is flexible enough to accommodate alternative weighting schemes in the analysis if desired.

However, DEA also suffers from some drawbacks. It can be vulnerable to data errors, because the DEA 'best practice' frontier is composed of a small number of highly performing organizations, and the performance of all other units is judged in relation to that frontier. Little can be said about organizations on the best practice frontier, as they are used as the basis for assessing the performance of all other organizations. Also, as more outputs (or more environmental – uncontrollable – factors) are included, an increasing number of organizations are likely to lie on the 'best practice' frontier, reducing the capacity to discriminate between organizations. This may be appropriate, but requires careful scrutiny. Finally, DEA measures technical efficiency and ignores allocative efficiency or overall cost-effectiveness. This means that an organization might be deemed efficient using DEA, but only if a zero weight is placed on an important output. This means that careful attention should be given to the 'slacks' on each input and output, as well as to the overall efficiency score.

Therefore, although DEA is a useful tool for exploring large and complex datasets and making preliminary comparisons, it is less well suited to testing hypotheses and drawing statistical inferences. Any DEA analysis should therefore examine a range of modelling perspectives in order to identify the sensitivity of judgements to different technical choices.

The Malmquist Index

Conventional DEA presents a cross-sectional analysis of efficiency. It is also possible to use DEA to examine secular trends in efficiency, and to decompose those trends into technological change and productivity growth. In this section we describe the use of a Malmquist index to explore productivity growth in one indicator of progress toward UHC, namely financial protection.

The three indices used most frequently to measure changes of productivity are the Törnqvist Index, Fisher's Ideal Index (the geometric mean of the Laspeyres and Paasche indices) and the Malmquist Productivity Index. The first two require the calculation of both the amounts and the prices of all inputs and outputs. In contrast, the Malmquist Index has the advantage that no information is needed on the prices of inputs and outputs. Furthermore, calculation of the Malmquist Index requires no restrictive assumptions regarding whether the organizations under analysis are benefit maximizers or cost minimizers. As Coelli, Rao and Battese (1998) indicate, these two characteristics make the Malmquist Index a particularly suitable tool for the analysis of productivity change in the public sector, where output prices are not in general available. A further advantage of the Malmquist approach is that it decomposes productivity into two parts that capture changes in the level of technical efficiency, and changes due to technical progress.

We use here the non-parametric methods of DEA to develop a series of production frontiers under different assumptions. The analytic framework can be illustrated graphically by means of Figure 5, which seeks to explain the Malmquist indices in intuitive form for a technology exhibiting variable returns to scale with just one input *x* and one output *y*. There are two time periods, *t* and *t*+1. The variable returns to scale (VRS) technology estimated by DEA in period *t* is represented by the frontier S^t_{VRS}, while the notional constant returns to scale (CRS) technology is indicated by the line S^t_{CRS}. The organization of interest consumes input x^t and produces output y^t in year *t*. Then we can examine the Malmquist Index as comprising three elements, M = (P x S) x T, as follows.

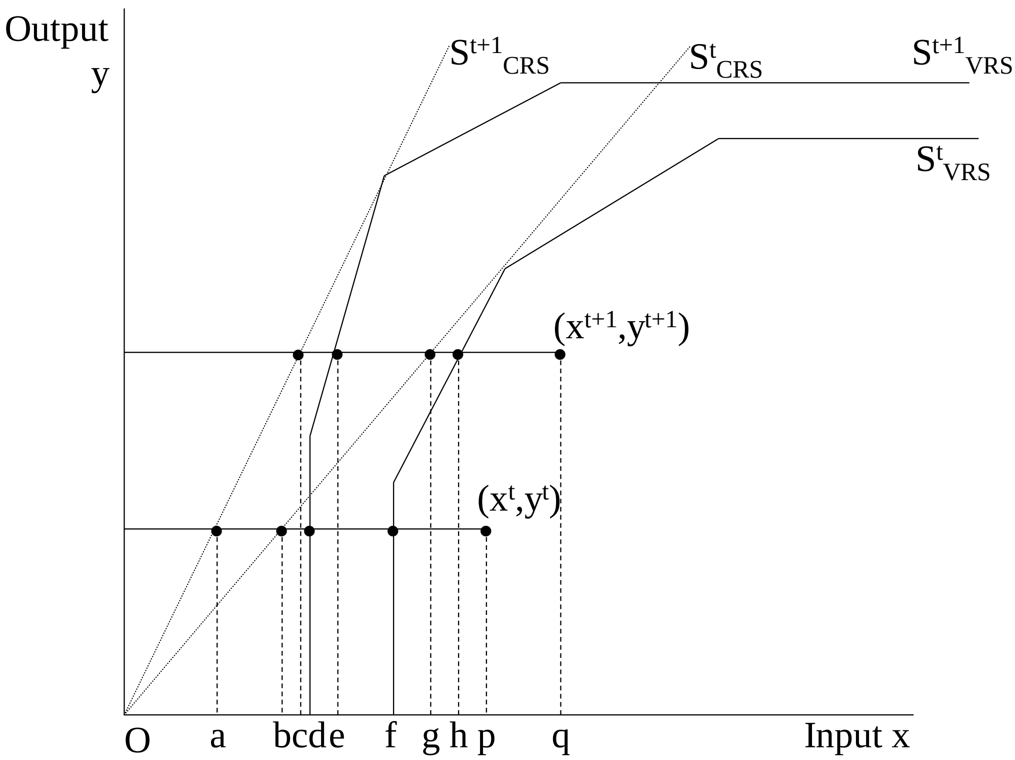

Figure 5: Illustration of productivity change with one input, one output

The pure efficiency change P between years is given by the ratio

$$P = \frac{(Oe/Oq)}{(Of/Op)}.$$

This simply indicates the change in the organization's distance from the current technically efficient frontier from one year to the next. The change in scale efficiency S is given by

$$S = \frac{\frac{(Oc/Oq)}{(Oe/Oq)}}{\frac{(Ob/Op)}{(Of/Op)}},$$

which indicates the change from one year to the next in the distance from the VRS frontier to the CRS frontier at the organization's observed level of inputs and outputs.

The change in the scale efficient technology indicated by the CRS frontiers is estimated by

$$\boldsymbol{T} = \sqrt{\left[\frac{(Og/Oq)}{(Oc/Oq)} . \frac{(Ob/Op)}{(Oa/Op)}\right]}.$$

This is the DEA estimate of the annual shift in the frontier under the assumption of constant returns to scale. (In this simple graphical example, the two components of T are identical, but this will not in general be the case.)

The Malmquist Index is then given by:

$$M = \frac{(Oe/Oq)}{(Of/Op)} . \left[\frac{\frac{(Oc/Oq)}{(Oe/Oq)}}{\frac{(Ob/Op)}{(Of/Op)}}\right] . \sqrt{\left[\frac{(Og/Oq)}{(Oc/Oq)} . \frac{(Ob/Op)}{(Oa/Op)}\right]}.$$

That is, the organization's productivity change is expressed as the product of pure efficiency change, scale efficiency change, and an estimate of technological progress.

An application of DEA to health coverage

Cross-sectional DEA applied to universal health coverage

We apply the DEA methods described above to examine the 'efficiency' with which countries achieve a given level of health coverage. The underlying model is as follows. The fundamental input required to secure coverage is 'pooled' spending on health services. These funds are deployed with varying levels of efficiency to provide access to needed health services and financial protection to citizens, the building blocks of 'universal coverage'. The analysis assesses the extent to which countries differ in the success with which their pooled funds secure a given level of coverage. A country's ability to maximize the impact of a given level of coverage on access may of course be affected by its levels of economic and social development. We therefore also examine the extent to which the results are altered by acknowledging that such factors may be uncontrollable influences on attainment, and including them as additional inputs in the analysis. We first develop a cross sectional model in this section, and in the following section undertake a Malmquist analysis.

In developing the models, we are highly constrained by data availability, especially as we wish to maintain consistency with the seven year panel data analysis presented later in the report. As a result we use the following model specification. We use as an input the *Pooled Spending Per Capita* (total health spending minus OOP) at PPP constant international $. This indicates the resources that a nation is directly devoting to financial protection and effective access in the health sector. We choose not to use just *publicly* pooled financing because the consequent coverage and access can arise regardless of who makes the prepayment. For example, some private prepayment arises from employer contributions. A separate study has examined whether public or private payment has more effect on health outcomes (Moreno-Serra and Smith, 2011).

As outputs we use three indicators of broader health system coverage: *Out of Pocket (OOP) payments* as percent of total health expenditure; *DTP3 immunization coverage* (% amongst 1-year-olds); and *Measles immunization* (% of children ages 12-23 months). Later we augment this analysis with further specifications in which *national income (GDP per capita at constant PPP$*) and *primary school education* are used as additional input constraints. DEA is highly sensitive to outlier observations, so we seek to minimize the influence of annual fluctuations by using average data for 2000-2006 for the 79 countries having a full data series available for that period.

Throughout, we use a modification of DEA proposed by Banker, Charnes and Cooper (1984) that constrains the weights on the efficient comparator DMUs to sum to one. This specification is required whenever (as in this case) data are in the form of ratios (eg rates *per capita*), rather than absolute numbers (Hollingsworth and Smith, 2003). Furthermore, we use throughout the 'input orientation' model specification. This indicates the extent to which pooled health spending could be reduced whilst still securing the same level of coverage. Results for the analogous 'output orientation' model are also available.

The results indicate the extent to which a country could secure the same level of coverage (in the form of protection from OOP expenditure, and measles and DTP3 immunization) with lower levels of pooled spending. It identified 18 efficient countries, as listed below in Table 1. The list also indicates the number of times the efficient country was cited as one of the efficient peers for another country. Note that the efficient high income countries were rarely used as peers for inefficient countries – they are likely to appear in the list because they have a uniquely high performance on one of the output dimensions. The more interesting comparators are therefore the efficient lower income countries that frequently act as peers for other countries, notably Mongolia, Eritrea, Kyrgyz Republic and Madagascar. The mean level of efficiency amongst the 79 countries was 0.442, indicating very high levels of performance amongst some low income countries, and large scope for improvement amongst some higher spending nations.

Country	Number of times a peer
Mongolia	43
Eritrea	26
Kyrgyz Republic	21
Madagascar	18
Oman	15
Gambia	13
Tanzania	11
Namibia	10
Guinea	7
Niger	7
Ethiopia	6
Egypt	5
Hungary	1
Tajikistan	1
Botswana	0
France	0
Luxembourg	0
Netherlands	0

Table 1: Efficient countries using basic DEA model, listed by number of times they appear as a peer for other countries

To illustrate the type of results available for each country, Table 2 shows the findings for Nicaragua, which had an efficiency level of 0.451. Its efficient peers were Tanzania (50.2%), Mongolia (16.9%), and Kyrgyz Republic (32.9%). The data used for the calculation of efficiency are shown below. The efficient composite has the same (or better) level of outputs as Nicaragua but uses only 0.451 of the inputs (pooled spending per capita of $42.1 rather than $93.2). Note that there is an additional slack for DTP3 immunization.

	Not OOP	DTP3	Measles	Pooled exp
Nicaragua	**58.3**	**84.6**	**93.6**	**93.2**
Tanzania	58.4	89.0	89.3	25.0
Mongolia	82.5	97.6	96.9	92.9
Kyrgyz Republic	45.9	97.6	98.4	42.0
Efficient composite*	**58.3**	**93.3**	**93.6**	**42.1**

* Comprises Tanzania (50.2%), Mongolia (16.9%) and Kyrgyz Republic (32.9%)

Table 2: Data for Nicaragua and its efficient peers using basic DEA model

It is of course always possible to argue with the choice of model, and DEA offers no tests of model specification with which to guide the choice of a preferred model. Countless variants of this basic model can be envisaged and tested. Ideally, if the relevant data were available, we would test a wider range of coverage indicators. However, given data constraints we limit further analysis to examining the effect of including as additional inputs certain influences on the levels of coverage achieved that are beyond the immediate control of the health system, and which therefore should be acknowledged when making comparisons between health systems.

By way of illustration, we test just two such variables: the level of primary school enrolment in the country, and its average income, as measured by GDP *per capita* in 2005 PPP$, which are beyond the control of the health system. In DEA, such uncontrollable factors are modelled by requiring that the organization can be compared only with a weighted combination of other organizations that suffers the same (or more adverse) level of that factor. Inclusion of an additional constraint will in general lead to increases in measured efficiency for many countries.

The inclusion of primary school enrolment has little impact on efficiency levels, suggesting that the frontier already contains many countries with low primary education levels. The exceptions are a handful of countries with very low levels of primary school enrolment (Burkina Faso, Djibouti and Swaziland) that experience large rises in efficiency levels when the variable is included as an input constraint. When income is included as a further input constraint, only Poland and Burundi experience significant efficiency increases. The mean level of efficiency amongst the 79 countries rises from 0.442 to 0.462, indicating that – after adjusting for such environmental factors – the scope for securing improved coverage performance is unchanged except amongst a small number of countries. In short, according to this analysis, adverse environmental circumstances cannot be used for an excuse for poor coverage levels except in a few cases. Full results are given in Annex 1.

Efficiency under the above analysis refers to the extent to which the input (expenditure) could be reduced whilst still maintaining achieved levels of coverage. In contrast, the output DEA orientation seeks to indicate the extent to which *outputs* could be increased using the same level of input. In the interests of brevity, this analysis is not covered in this paper. The analysis yields a similar ordering of countries, but less dispersion of efficiency levels, with an average level of efficiency of 94.7%. This is because – to get a low score – inefficient countries must demonstrate scope for improvement across all three outputs. In other words, they can avoid low efficiency scores by securing good performance in only one output dimension. Some countries nevertheless indicate considerable scope for improved overall coverage according to this demanding criterion, including Lao PDR (0.617), India (0.654), Azerbaijan (0.762), Togo (0.764) and Ecuador (0.775).

Explaining efficiency variations

A further potential analytic step is to seek to explain statistically the DEA efficiency scores. There is some controversy in the productivity literature on whether this can be theoretically justified, and – if so – what methods to use (Jacobs, Smith and Street, 2006). In this paper, to the extent that data permit, we merely explore associations between the DEA scores and the characteristics of the health systems. It is important to note that these characteristics should be policy choices related to the organization and governance of the health system, and not exogenous determinants of outcomes that should have been captured in the initial DEA analysis.

There is a severe shortage of relevant data that can be used for such purposes. We use the variables set out in Table 3, which can be considered under two broad headings: organization of health care resources, and general public governance. Variables related to the organization of health resources are drawn from the World Health Organization's *Global Health Observatory* database (http://www.who.int/gho/database/en), whereas the source for governance indicators is the World Bank's *Worldwide Governance Indicators (WGI)* database (http://www.govindicators.org). As in the DEA work above, country figures refer to pooled averages for years 2000-2006.

SHI_share	Social security expenditure as a percentage of government health expenditure
Doctors_pc	Physicians (per 1,000 people)
Nurses_pc	Nurses and midwives (per 1,000 people)
Beds_pc	Hospital beds (per 1,000 people)
Governance_mean	WGI governance index: average of six dimensions (high = better)
No_corruption	Control of Corruption WGI index
Effective_govt	Government Effectiveness WGI index
Stability	Political Stability and Absence of Violence/Terrorism WGI index
Reg_quality	Regulatory Quality WGI index
Rule_of_law	Rule of Law WGI index
Voice_Accountab	Voice and Accountability WGI index

Table 3: Variables used to explain efficiency scores

DEA efficiency scores are censored with an upper limit of 1.0. Therefore it is conventional to model them using a censored (tobit) regression. In this application we present three such regressions, as follows:

1. A model with only health services organizational variables
2. A model with health services organizational variables plus a composite governance variable
3. A model with individual governance indicators.

We used DEA scores from the model with education and income adjustments. The results are set out in Table 4. They suggest that the intensity of doctors employed in the health system is consistently associated with lower DEA efficiency measures. They also suggest - apparently perversely – that a higher average index of governance is associated with lower DEA efficiency. However, the latter index is an average constructed from many other average indicators of specific governance dimensions, thus potentially introducing a considerable amount of 'noise' into the regressions. We therefore also estimate the model using all the available specific governance indicators separately. This regression suggests a more nuanced pattern, with effective government and voice/accountability having a negative association whilst rule of law is positively associated with DEA scores.

It is possible to argue that a higher reliance on doctors to deliver basic health services might lead to higher costs, and therefore lower measured efficiency levels. However, it is hard to see why countries with higher efficiency levels could be associated with poorer performance in particular governance dimensions. It is most likely that poor governance is associated with some other factor (such as high NGO penetration) that has led to better coverage in the domains we are able to measure. However, we urge extreme caution in interpreting these results. They can be no more than suggestive of future research possibilities in a new and complex area of health systems research.

	Model 1	Model 2	Model 3
SHI_share	0.005	0.003	0.019
	(0.021)	(0.021)	(0.020)
Doctors_pc	-0.227***	-0.204***	-0.150**
	(0.064)	(0.064)	(0.059)
Nurses_pc	0.001	0.025	0.026
	(0.019)	(0.021)	(0.020)
Beds_pc	0.035	0.037	0.012
	(0.038)	(0.036)	(0.033)
Governance_mean		-0.191**	
		(0.077)	
No_corruption			0.004
			(0.196)
Effective_govt			-0.708**
			(0.270)
Stability			0.094
			(0.086)
Reg_quality			0.037
			(0.214)
Rule_of_law			0.624**
			(0.245)
Voice_Accountab			-0.200*
			(0.105)

Table 4: Tobit regressions of health system characteristics on DEA efficiency scores
Notes: Results from cross-sectional regressions using country averages for years 2000-2006. All models estimated with standard errors (in parentheses under coefficients) robust to arbitrary heteroskedasticity. * Significant at 10%; ** significant at 5%; *** significant at 1%.

An application of the Malmquist Index to universal health coverage

In this section we use the same indicators of universal coverage introduced in the cross-sectional DEA section to explore productivity growth in UHC amongst the 79 countries, having a full panel of data from 2000 to 2006. As above, data limitations severely constrain the size of the panel, the chosen indicators, and the time period under scrutiny. The results use the simplest model specification, with one input (pooled health care spending) and three outputs (OOP protection, and measles and DTP3 immunization). Because the data are expressed as ratios, constant returns to scale are already assumed, and we can make no comment on scale efficiency. We therefore report only on technological and efficiency change.

The results are summarized in the Table 5, which indicates the proportionate change in (a) efficiency (b) technology and (c) total factor productivity in the sample in each year. Thus, over the full seven year period 2000-2006 (the bottom row), total productivity fell at a rate of 3.9% per annum (reflecting the annual growth factor 0.961), the net effect of a 5.9% per annum decrease in efficiency being balanced to some extent by a 2.1% per annum advance in the frontier (technological advance). From year to year there is considerable volatility in efficiency and technological change, probably brought about by fluctuations in the identity and performance of the outliers that form the frontier in each year. The broad conclusion is that – whilst the frontier has advanced somewhat – the dispersion of inefficient countries has increased, giving rise to a greater average deviation from the efficient frontier. Annex 2 gives the average annual changes in efficiency and technology for each of the 79 countries in the sample.

	Efficiency growth	**Technology growth**	**Total productivity growth**
2001	0.969	0.983	0.952
2002	1.008	0.954	0.962
2003	0.899	1.066	0.959
2004	0.848	1.144	0.971
2005	1.012	0.949	0.960
2006	0.919	1.046	0.962
mean	**0.941**	**1.021**	**0.961**

Table 5: Mean productivity growth factors 2001 to 2006 (1.000 = no change)

Because of the volatility in the annual changes, the policy interpretation of this analysis must be somewhat cautious. It does suggest that there has been an advance in the 'best practice' frontier over the seven year period, in the sense that the best health systems are able to offer increasing levels of access and financial protection for given levels of pooled financing. However, the decrease in efficiency over the period suggests that some countries, such as Moldova and Tanzania, are failing to keep pace with that best practice, and are falling further behind.

Discussion

DEA offers a useful device for exploring outlying performance and the dispersion of efficiency levels. The analysis in this paper shows the extent to which countries are using pooled sources of health system finance to secure effective coverage and access for their populations. The results must be interpreted with great caution, given the scope for endless debate about the precise specification of the model to be used. However, they do suggest wide variations in performance.

There is no formal test for DEA model specification and results, although the successive incorporation of various putative 'uncontrollable' influences on performance allows us to scrutinize the reasons why a country secures a favourable ranking. Our exploratory analysis of the association between health system characteristics and DEA efficiency scores yields some apparently perverse results that deserve further study, but require more detailed data before further analysis can be undertaken.

One of the benefits of DEA is that it facilitates detailed scrutiny of individual countries' performance, as in the example of Nicaragua given above. This can be helpful for decision-makers wishing to understand where the major scope for improvement lies in their country, and also what the relevant 'best practice' peers might be. Presentation in full of such detailed benchmarking data is infeasible. However, Annex 3 gives a summary of the data from the basic DEA model to give an indication of the type of material that can be made available.

Although offering some promise in theory to address the questions regarding the efficiency with which progress towards UHC is advancing, the Malmquist analysis has uncovered great volatility in estimates of technological progress and efficiency change, as is often the case. This is probably due to the sensitivity of DEA to stochastic variations in the outlying best practice countries. Over the entire seven year period under scrutiny, it is likely that the estimates of annual technological progress (growth of 2.1% per annum) and efficiency change (annual decline of 5.9%) are informative of important qualitative changes in system performance. However, it would be prudent not to read too much into these figures. Again, further progress will be seriously limited by data availability.

More generally, we underline the value of DEA as an exploratory tool rather than offering a definitive judgement on health system performance. If used in that spirit, it can offer useful diagnostic information. We have presented a small number of models, and there is a more detailed Excel spreadsheet of results that can be scrutinized.

References

Anand, S. *et al*, (2002), *Report of the Scientific Peer Review Group on health systems performance assessment*, Geneva: World Health Organization (pp 141).

Andersen, P. and Petersen, N. C. (1993), "A procedure for ranking efficient units in Data Envelopment Analysis", *Management Science*, 39, 1261-1264.

Banker, R.D., A.Charnes and W.W. Cooper (1984) 'Models for estimation of technical and scale inefficiencies in data envelopment analysis' *Management Science*, 30, 1078 1092.

Charnes, A., W.W. Cooper and E. Rhodes (1978), 'Measuring the efficiency of decision making units', *European Journal of Operational Research*, 3, 429-444.

Charnes, A., Cooper, W., Lewin, A. Y. and Seiford, L. M. (1994) (eds), *Data envelopment analysis: theory, methodology and applications*, Boston: Kluwer Publishers.

Coelli, T., Rao, D.S.P. and Battese, G.E. (1998), *An Introduction to Efficiency and Productivity Analysis*. Kluwer, Dordrecht.

Emrouznejad, A., Parker, B. R. and Tavares, G. (2008), "Evaluation of research in efficiency and productivity: A survey and analysis of the first 30 years of scholarly literature in DEA", *Socio-Economic Planning Sciences*, 42(3), 151-157.

Farrell, M.J. (1957), "The measurement of productive efficiency", *Journal of the Royal Statistical Society, Series A*, 120, 253-281.

Hollingsworth, B. and Smith, P. (2003), "The use of ratios in data envelopment analysis", *Applied Economics Letters*, 10(11), 733-735.

Jacobs, R., Smith, P. and Street, A. (2006), *Measuring efficiency in health care: analytic techniques and health policy*, Cambridge: Cambridge University Press.

Moreno-Serra, R. and Smith, P. C. (2011). The effects of health coverage on population outcomes: a country-level panel data analysis. Results for Development Institute Working Paper. Washington: Results for Development Institute.

Smith, P. and Street, A. (2005), "Measuring the efficiency of public services: the limits of analysis", *Journal of the Royal Statistical Society, Series A*, 168(2), 401-417.

World Health Organization (2000), *World Health Report 2000*, Geneva: WHO.

Annex 1: Data envelopment analysis results for 2000-2006 pooled data: input and output oriented formulations

This annex presents cross-sectional DEA results for 79 countries for the pooled average of years 2000-2006. The 'basic' model uses the three health coverage measures as outputs and pooled health care finance (public plus private) as input. To this is then added the percentage of children of primary school age enrolled in education as a fixed input constraint (only comparisons with the same or lower enrolment levels will be considered). Finally, GDP per capita is added as another fixed input constraint (only comparisons with the same or lower income levels will be considered). The results are presented in both input- and output-oriented format.

Annex 1: DEA results, all countries

Country	DEA Efficiency		
	Basic model	Add primary education	Add income & education
Algeria	0.304	0.304	0.304
Australia	0.038	0.038	0.038
Azerbaijan	0.216	0.216	0.216
Bahrain	0.926	0.926	0.926
Belgium	0.034	0.034	0.034
Belize	0.354	0.354	0.354
Botswana	1.000	1.000	1.000
Bulgaria	0.117	0.117	0.117
Burkina Faso	0.477	0.690	0.690
Burundi	0.929	0.929	1.000
Cape Verde	0.704	0.704	0.704
Cyprus	0.074	0.074	0.074
Denmark	0.077	0.077	0.077
Djibouti	0.263	1.000	1.000
Dominican Rep	0.085	0.085	0.085
Ecuador	0.094	0.094	0.094
Egypt	1.000	1.000	1.000
Eritrea	1.000	1.000	1.000
Estonia	0.145	0.145	0.145
Ethiopia	1.000	1.000	1.000
Fiji	0.508	0.508	0.508
Finland	0.118	0.118	0.118
France	1.000	1.000	1.000
Gambia	1.000	1.000	1.000
Germany	0.128	0.128	0.128
Ghana	0.382	0.382	0.382
Greece	0.037	0.037	0.037
Guatemala	0.206	0.206	0.206
Guinea	1.000	1.000	1.000
Hungary	1.000	1.000	1.000
Iceland	0.035	0.035	0.035
India	0.390	0.390	0.390
Indonesia	0.453	0.453	0.453
Ireland	0.077	0.077	0.077
Israel	0.055	0.055	0.055
Italy	0.038	0.038	0.038
Japan	0.204	0.204	0.204
Kazakhstan	0.524	0.524	0.524
Korea, Rep.	0.087	0.087	0.087
Kuwait	0.426	0.426	0.426

Country	DEA Efficiency Basic model	Add primary education	Add income & education
Kyrgyz Republic	1.000	1.000	1.000
Lao PDR	0.463	0.463	0.463
Lesotho	0.636	0.636	0.636
Lithuania	0.204	0.204	0.204
Luxembourg	1.000	1.000	1.000
Macedonia, FYR	0.173	0.173	0.173
Madagascar	1.000	1.000	1.000
Malaysia	0.185	0.185	0.185
Mauritania	0.666	0.666	0.666
Mexico	0.135	0.135	0.135
Moldova	0.649	0.649	0.649
Mongolia	1.000	1.000	1.000
Morocco	0.420	0.420	0.420
Namibia	1.000	1.000	1.000
Netherlands	1.000	1.000	1.000
New Zealand	0.061	0.061	0.061
Nicaragua	0.451	0.451	0.451
Niger	1.000	1.000	1.000
Norway	0.033	0.033	0.033
Oman	1.000	1.000	1.000
Panama	0.150	0.150	0.150
Peru	0.299	0.299	0.299
Poland	0.767	0.767	1.000
Romania	0.379	0.379	0.379
Slovenia	0.239	0.239	0.239
South Africa	0.148	0.148	0.148
Spain	0.052	0.052	0.052
Swaziland	0.638	1.000	1.000
Sweden	0.188	0.188	0.188
Switzerland	0.019	0.019	0.019
Tajikistan	1.000	1.000	1.000
Tanzania	1.000	1.000	1.000
Togo	0.755	0.755	0.755
Tunisia	0.274	0.274	0.274
Turkey	0.149	0.149	0.149
United Arab Emirates	0.073	0.077	0.077
United Kingdom	0.086	0.086	0.086
USA	0.047	0.047	0.047
Venezuela, RB	0.055	0.055	0.055
Mean	**0.442**	**0.458**	**0.462**

Annex 2: Annual means for countries in Malmquist analysis

Country	Δ Efficiency	Δ Technology	Δ Total
Algeria	0.929	1.011	0.939
Australia	0.961	1.012	0.973
Azerbaijan	0.859	1.049	0.901
Bahrain	0.959	1.018	0.976
Belgium	0.966	1.013	0.979
Belize	0.950	1.045	0.992
Botswana	0.906	1.007	0.912
Bulgaria	0.887	1.031	0.915
Burkina Faso	0.925	1.019	0.942
Burundi	0.886	1.026	0.909
Cape Verde	0.937	1.015	0.951
Cyprus	0.909	1.044	0.949
Denmark	0.950	1.011	0.961
Djibouti	0.977	1.000	0.978
Dominican Republic	0.957	1.035	0.990
Ecuador	0.884	1.026	0.907
Egypt	0.909	1.049	0.953
Eritrea	1.049	1.057	1.109
Estonia	0.911	1.015	0.924
Ethiopia	0.974	1.003	0.977
Fiji	1.011	1.014	1.025
Finland	0.936	1.016	0.951
France	0.969	1.007	0.976
Gambia	0.941	1.013	0.953
Germany	0.979	1.009	0.989
Ghana	0.962	1.035	0.996
Greece	0.911	1.032	0.940
Guatemala	0.980	1.055	1.034
Guinea	0.962	1.028	0.988
Hungary	0.914	1.020	0.932
Iceland	0.967	1.013	0.979
India	0.918	1.058	0.971
Indonesia	0.952	1.022	0.974
Ireland	0.944	1.005	0.950
Israel	0.966	1.023	0.988
Italy	0.967	1.013	0.979
Japan	0.967	1.013	0.979
Kazakhstan	0.890	1.037	0.922
Korea, Rep.	0.885	1.036	0.916
Kuwait	0.996	1.016	1.013
Kyrgyz Republic	0.873	1.050	0.917
Lao PDR	0.902	1.017	0.918

Country	Δ Efficiency	Δ Technology	Δ Total
Lesotho	0.963	1.017	0.980
Lithuania	0.912	1.023	0.933
Luxembourg	0.932	1.007	0.939
Macedonia, FYR	0.962	1.022	0.984
Madagascar	1.002	0.999	1.001
Malaysia	0.902	1.027	0.927
Mauritania	1.003	1.005	1.008
Mexico	0.922	1.055	0.972
Moldova	0.816	1.042	0.850
Mongolia	0.976	1.014	0.990
Morocco	0.891	1.051	0.936
Namibia	0.944	0.999	0.943
Netherlands	0.963	1.007	0.970
New Zealand	0.943	1.007	0.950
Nicaragua	0.931	1.033	0.961
Niger	0.923	0.990	0.913
Norway	0.973	1.008	0.980
Oman	1.012	1.011	1.022
Panama	0.968	1.018	0.985
Peru	0.955	1.023	0.977
Poland	0.925	1.022	0.945
Romania	0.934	1.019	0.953
Slovenia	0.954	1.009	0.963
South Africa	0.970	0.998	0.967
Spain	0.942	1.017	0.958
Swaziland	0.951	1.010	0.961
Sweden	0.948	1.012	0.959
Switzerland	0.964	1.019	0.982
Tajikistan	0.893	1.008	0.900
Tanzania	0.856	1.039	0.890
Togo	0.939	1.058	0.993
Tunisia	0.935	1.032	0.965
Turkey	0.955	1.009	0.964
United Arab Emirates	1.000	1.014	1.013
United Kingdom	0.942	1.007	0.948
United States of America	0.953	1.009	0.962
Venezuela, RB	0.940	1.046	0.983

Annex 3: Output for basic model, input orientation

This table presents the results of the analysis for the basic three-output, one-input model under input orientation. The efficiency score shows the extent to which the input (pooled resources) could be reduced whilst keeping outputs constant. The table shows the number of times each country is used as a peer for other countries. It then shows the peers that were chosen to yield the country's efficiency score (when that score is 1.000, the only peer is the country itself). The final columns show the weights attached to each of the peers in creating the efficient composite benchmark for the country.

	Country	Efficiency	Used as peer	Peers (numbers refer to left hand column)				Weights on peers			
1	Algeria	0.304	0	24	52	47		0.461	0.328	0.211	
2	Australia	0.038	0	24	52	47		0.123	0.811	0.066	
3	Azerbaijan	0.216	0	18	29	58		0.533	0.420	0.048	
4	Bahrain	0.926	0	41	60	30		0.058	0.294	0.647	
5	Belgium	0.034	0	24	18	52		0.195	0.072	0.733	
6	Belize	0.354	0	18	41	72	52	0.033	0.475	0.315	0.178
7	Botswana	1.000	0	7				1.000			
8	Bulgaria	0.117	0	18	41	52	72	0.416	0.145	0.324	0.115
9	Burkina Faso	0.477	0	18	58	29		0.533	0.434	0.034	
10	Burundi	0.929	0	47	18	20		0.180	0.707	0.113	
11	Cape Verde	0.704	0	24	52	47		0.638	0.359	0.003	
12	Cyprus	0.074	0	41	52			0.903	0.097		
13	Denmark	0.077	0	60	52	54		0.208	0.720	0.072	
14	Djibouti	0.263	0	47	18	20		0.310	0.108	0.582	
15	Dominican Republic	0.085	0	47	18	20		0.108	0.803	0.089	
16	Ecuador	0.094	0	18	58	29		0.644	0.249	0.107	
17	Egypt	1.000	5	17				1.000			
18	Eritrea	1.000	26	18				1.000			
19	Estonia	0.145	0	24	52	72		0.086	0.808	0.106	
20	Ethiopia	1.000	6	20				1.000			
21	Fiji	0.508	0	24	18	52		0.487	0.064	0.450	
22	Finland	0.118	0	52	60	17		0.565	0.304	0.131	
23	France	1.000	0	23				1.000			
24	Gambia	1.000	13	24				1.000			
25	Germany	0.128	0	60	52	54		0.627	0.278	0.095	
26	Ghana	0.382	0	47	18	20		0.035	0.809	0.156	
27	Greece	0.037	0	18	41	72	52	0.043	0.224	0.442	0.290
28	Guatemala	0.206	0	18	41	71		0.537	0.269	0.194	

	Country	Efficiency	Used as peer	Peers (numbers refer to left hand column)				Weights on peers			
29	Guinea	1.000	7	29				1.000			
30	Hungary	1.000	1	30				1.000			
31	Iceland	0.035	0	24	52	47		0.022	0.935	0.043	
32	India	0.390	0	18	58	29		0.286	0.133	0.580	
33	Indonesia	0.453	0	47	18	20		0.155	0.554	0.291	
34	Ireland	0.077	0	52	54	47		0.470	0.394	0.135	
35	Israel	0.055	0	72	52	41		0.146	0.602	0.251	
36	Italy	0.038	0	24	52	47		0.430	0.532	0.039	
37	Japan	0.204	0	60	41	52		0.779	0.090	0.131	
38	Kazakhstan	0.524	0	52	60	41		0.260	0.067	0.672	
39	Korea, Rep.	0.087	0	52	72	41		0.354	0.049	0.598	
40	Kuwait	0.426	0	52	60	17		0.172	0.621	0.207	
41	Kyrgyz Republic	1.000	21	41				1.000			
42	Lao PDR	0.463	0	18	58	29		0.093	0.557	0.351	
43	Lesotho	0.636	0	24	18	72	47	0.490	0.177	0.085	0.247
44	Lithuania	0.204	0	52	60	41		0.621	0.082	0.297	
45	Luxembourg	1.000	0	45				1.000			
46	Macedonia, FYR	0.173	0	52	72	41		0.615	0.207	0.178	
47	Madagascar	1.000	18	47				1.000			
48	Malaysia	0.185	0	18	41	52		0.500	0.040	0.460	
49	Mauritania	0.666	0	47	18	20		0.389	0.265	0.347	
50	Mexico	0.135	0	52	41			0.057	0.943		
51	Moldova	0.649	0	18	41	52		0.088	0.740	0.172	
52	Mongolia	1.000	43	52				1.000			
53	Morocco	0.420	0	41	18			0.559	0.441		
54	Namibia	1.000	10	54				1.000			
55	Netherlands	1.000	0	55				1.000			
56	New Zealand	0.061	0	52	54	47		0.678	0.105	0.217	

	Country	Efficiency	Used as peer	Peers (numbers refer to left hand column)				Weights on peers			
57	Nicaragua	0.451	0	72	52	41		0.502	0.169	0.329	
58	Niger	1.000	7	58				1.000			
59	Norway	0.033	0	54	52	47		0.139	0.753	0.108	
60	Oman	1.000	15	60				1.000			
61	Panama	0.150	0	24	18	52		0.018	0.239	0.743	
62	Peru	0.299	0	41	52	72		0.088	0.403	0.509	
63	Poland	0.767	0	17	60			0.143	0.857		
64	Romania	0.379	0	52	41	60	17	0.681	0.028	0.060	0.231
65	Slovenia	0.239	0	60	52	54		0.673	0.211	0.116	
66	South Africa	0.148	0	54	52	47		0.184	0.130	0.686	
67	Spain	0.052	0	72	52	41		0.060	0.830	0.109	
68	Swaziland	0.638	0	60	52	54		0.034	0.810	0.156	
69	Sweden	0.188	0	60	17	52		0.869	0.091	0.040	
70	Switzerland	0.019	0	24	18	52		0.538	0.254	0.208	
71	Tajikistan	1.000	1	71				1.000			
72	Tanzania	1.000	11	72				1.000			
73	Togo	0.755	0	18	29	58		0.463	0.429	0.108	
74	Tunisia	0.274	0	18	41	52		0.088	0.499	0.413	
75	Turkey	0.149	0	52	24	47		0.425	0.426	0.149	
76	United Arab Emirates	0.073	0	24	52	47		0.129	0.818	0.053	
77	United Kingdom	0.086	0	52	60	54		0.622	0.033	0.344	
78	USA	0.047	0	60	52	54		0.285	0.539	0.176	
79	Venezuela, RB	0.055	0	58	18	29		0.129	0.771	0.100	